

PCT

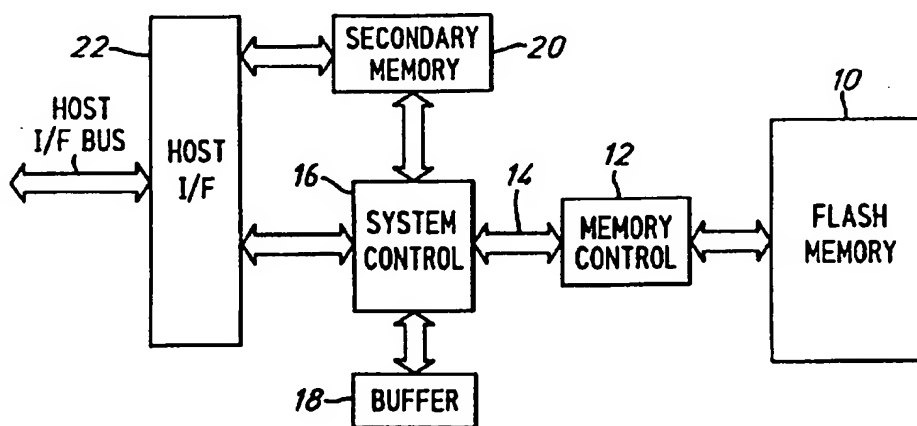
WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G06F 12/02, 3/06, G11C 16/06	A1	(11) International Publication Number: WO 97/12325 (43) International Publication Date: 3 April 1997 (03.04.97)
(21) International Application Number: PCT/GB96/00250 (22) International Filing Date: 6 February 1996 (06.02.96) (30) Priority Data: 9519670.5 27 September 1995 (27.09.95) GB (71) Applicant (for all designated States except US): MEMORY CORPORATION PLC [GB/GB]; The Computer House, Dalkeith Palace, Dalkeith, Edinburgh EH22 2NA (GB). (72) Inventor; and (75) Inventor/Applicant (for US only): SINCLAIR, Alan, Welsh [GB/GB]; 102 Lamb's Lane, Cottenham, Cambridge CB4 4TA (GB). (74) Agent: CULLIS, Roger; British Technology Group Ltd., 101 Newington Causeway, London SE1 6BU (GB).		(81) Designated States: CN, JP, KR, SG, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>

(54) Title: **MEMORY SYSTEMS**



(57) Abstract

A solid state memory for emulating a disk drive comprising: translation means for translating a logical sector address to a main memory address; a main memory composed of non-volatile memory cells erasable in blocks; characterised in that a first pointer is used to point to an unwritten location in main memory, and a second pointer is used to point to the next unerased erasable block in sequence to the erasable block containing the said unwritten memory location; control means being provided to ensure that there is always at least one erasable block in the erased condition between the first and second pointers.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

Memory Systems

This invention relates to memory systems and, in particular, to memory systems for storing block structured data. It finds application where non-volatile memory cells are used as the storage medium. The particular type of non-volatile memory cells used in this invention are memory cells which do not allow erasure of individual cells but allow erasure of blocks of memory cells. Different memory cells, such as FLASH EPROM cells, chalcogenide memory cells, and ferro-optic cells may be used.

Many common memory storage devices for computers are based on rotating storage media such as the magnetic disk drive. Rotating storage media have, however, a few disadvantages. They require comparatively large amounts of power because they consist of rotating platters. They are comparatively slow to read from and write to because of the seek latency (the time taken to move the head to the correct location) and the rotational latency (the time taken to locate the correct part of a track). Magnetic disks are also comparatively heavy and not very robust (they are sensitive to being knocked). In contrast, semiconductor memory is lightweight, robust, consumes very little power, and is fast in operation. For these reasons it is desirable to use a solid state equivalent of the magnetic disk. The particular semiconductor chosen as the solid state memory is preferably high density, non-volatile, and has low cost per bit. This makes FLASH memory an ideal candidate.

There are many prior art references to the use of FLASH memory as the storage medium for a solid state disk drive. These references include the following patents: EP 0 392 895, EP 0 424 191, EP 0 522 780, EP 0 597 706, EP 0 617 363, EP 0 618 535, EP 0 619 541, GB 2 251 323, WO 93 02418, WO 94 19746, WO 94 20906, WO 94 23369, WO 94 23432,

WO 95 14272, WO 95 10083, US 4 511 964, WO 84 00628, WO 95 10083, WO 95 14272.

Many of these references have certain features in common. For example, most references disclose the use of a table for converting the logical address from the host computer (usually in CHS - cylinder, head and sector-format) to a physical address in the FLASH

5 memory or to an intermediate address which will be used for generating the physical address in FLASH memory. Most references also have some means for ensuring that one block is not written to substantially more times than another block. A third common feature in FLASH solid state disk emulators is a buffer for increasing the write speed to the semiconductor. The buffer is usually made of some fast memory such as SRAM (Static
10 Random Access Memory).

Any solid state disk emulation system which uses FLASH memory as the storage medium must also have some method of controlling the erasure and writing of the cells. The reason for this is that the cells take a comparatively long time to erase and a cell can only be written to once before it needs to be erased. The present invention uses an
15 algorithm which may be used in a solid state disk emulating system for controlling the writing and erasure of memory cells.

EP 0 522 780 describes a control method for a solid state disk which uses a memory managing table recording the number of times each cell has been erased, the status of each memory block and a buffer to hold data which is to be written to memory. The method
20 involves copying sectors which are to be erased to another block before erasing the initial block.

EP 0 392 895 relates to the use of a cache memory for buffering writes to the FLASH memory to increase the speed of operation of the memory. The memory system

in that disclosure uses means to determine "the time since each data file was last written". It also relates to correcting for errors and defect substitution, but these were the subject of divisional applications which are discussed below.

EP 0 618 535 relates to a solid state disk emulator which has defect substitution.

5 EP 0 617 363 relates to a method of correcting for defective cells in a FLASH memory by using pointers to point from the defective locations in the memory to the spare locations.

EP 0 424 191 relates to a method of storing sequential data in a memory which is accessible in a prescribed order. Defects in the memory are stored in an ordered directory
10 so that these defective locations are skipped when the memory is being written to.

EP 0 619 541 provides a solid state disk emulating system which allows random access from a central processing unit. It also measures the erasure count for each block of memory so that a block can be avoided if it has been written to a large number of times.

EP 0 597 706 discloses a solid state peripheral storage device which uses a MAP
15 ROM to map the logical sector address provided by the host to a physical address in the semiconductor memory. The addresses of any bad sectors in semiconductor memory are recorded in the MAP ROM. A microsequencer is used to control the mapping.

GB 2 251 323 describes a disk emulation system for FLASH memory. The clean-up operation described in that disclosure is only automatic in the sense that it happens as
20 a background task when certain conditions are met. Either the least cycled block is erased or the block with the highest number of deleted sectors is erased. This gives rise to an essentially random block selection for erasure.

WO 94 20906 discloses a FLASH file system which is used to emulate the

operation of a disk drive. It uses a block allocation map to store information on the blocks in FLASH memory that can be written to. When a block is to be updated the allocation map is scanned until a free block is located. A transfer unit is used to facilitate memory reclaiming of the system. The transfer unit is an unwritten erased block of memory. The
5 active (currently valid) data from a block of memory that contains old data (the old block) is written to the same locations in the transfer unit (the new block) and the old block is erased. The old block then becomes the transfer unit.

WO 94 23432 discloses a FLASH memory mass storage architecture. Erase cycles are avoided by programming altered data file into an empty mass storage block.
10 Periodically the storage is cleaned up with a multi-sector erase. A counter is used to ensure that no block is erased more than a maximum number of times.

WO 94 23369 discloses a FLASH memory mass storage architecture which is very similar to that disclosed in WO 94 23432. Both disclosures are by the same applicant.

WO 94 19746 discloses a FLASH solid state drive which stores two bytes at a time.
15 Each time a sector is written into FLASH memory a header is also written with it. If an older version of that sector exists in the memory then it is marked as invalid. When storage space dips below a certain threshold then a clean up operation is performed. Good data from a sector is copied to a buffer, the sector is erased and the good data is copied back to the sector.

20 WO 84 00628 discloses a method of managing defects in a storage media. The method involves receiving an address from a host, compensating the received address for the number of defective locations below that address, and accessing the compensated address. A sequentially ordered list of defective locations in the storage media is

constructed.

WO 95 10083 uses counters to count the number of erase cycles each block has experienced and copies updated versions of data files to new locations. Periodically, blocks are erased to clean up the system. The blocks in the semiconductor memory are sized to
5 conform to commercial hard disk sector sizes.

WO 95 14272 provides a method of generating the correct physical starting address from the number of the logical block to be accessed.

A preferred embodiment of the invention provides automatic equalisation of erase/write cycles throughout the entire memory. This automatic equalisation is automatic
10 in the fullest sense in that the equalisation occurs as data is being written, rather than as a background task at some later time. It may also provide automatic compaction of the stored data. It may also achieve 100% efficiency in the use of available FLASH memory capacity, that is, all available FLASH memory is used. It may also permit the use of FLASH memory with any erasable block size: the erasable block size does not have to be 512 bytes.
15 This permits error correction codes to be included with each segment of data stored. The invention may also facilitate a very simple translation from physical to FLASH address space.

The present invention provides a solid state memory for emulating a disk drive comprising: translation means for translating a logical address to a second address for
20 accessing a store; a store composed of non-volatile memory cells erasable in blocks; characterised in that a first pointer is used to point to an unwritten location in the store, and a second pointer is used to point to the next unerased erasable block in sequence to the erasable block containing the said unwritten memory location; control means being

provided to ensure that there is always at least one erasable block in the erased condition between the first and second pointers. A solid state memory where the said unwritten location in the store is the lowest unwritten location in the store.

A solid state memory where the said unwritten location in the store is the highest
5 unwritten location in the store.

A solid state memory where the said next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is the unerased erasable block with the lowest physical address above the first pointer.

A solid state memory where the said next unerased erasable block in sequence to
10 the erasable block containing the said unwritten memory location is the unerased erasable block with the highest physical address below the first pointer.

A solid state memory, where the said next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is determined by a suitable algorithm. A solid state memory where a buffer is used to increase the speed of
15 writing data to the store.

A solid state memory where the said buffer is an SRAM buffer.

A solid state memory where the said buffer is a DRAM buffer.

A solid state memory, where the said store is FLASH EPROM.

A solid state memory where the said the store is composed of chalcogenide cells.

20 A solid state memory where the said the store is a ferro-optic memory.

A solid state memory where the said control means are provided by a microprocessor, a microcontroller, or a state machine.

In the present application an unwritten memory location does not refer to a memory

location which has never been written to, but rather it refers to a location which is in the erased condition.

In the present application a second address may be a physical address, a physical block address, or an intermediate address.

5 The invention will be particularly described, by way of example, with reference to the accompanying drawings, in which:

Figure 1 shows a diagram of a memory system for emulating a magnetic disk drive;

Figure 2 shows the hierarchy of subdivisions within a FLASH memory;

10 Figures 3A and 3B illustrate how the write/erase algorithm manages the location and frequency of erase and write operations in main FLASH memory.

Figure 3A shows a FLASH memory prior to the erase/write algorithm taking effect. Figure 3B shows the same FLASH memory after the write/erase algorithm has taken effect.

Referring to Figure 1, a FLASH memory 10 is used as the main store for data
15 received from the host. The FLASH memory 10 is controlled by a memory controller 12. The memory controller 12 operates between the FLASH memory 10 and a technology independent interface 14. The memory controller 12 controls the low-level execution of high level functions within the physical memory (FLASH memory 10). The memory controller 12 is memory technology dependent.

20 The technology independent interface 14 connects the memory controller 12 to a system controller 16. The system controller 16 manages the storage and accessing of block-structured data in the main FLASH memory 10 and in a write buffer 18 which is included to increase the speed of writing data to FLASH memory 10. The system

controller 16 performs the following functions. It translates the logical address from the host to an address suitable for use with the FLASH memory 10, hereinafter referred to as the logical to physical address translation. This translation may result in one or more intermediate addresses being generated. It maintains any address lookup tables which are
5 needed. It manages defective storage locations in the FLASH memory 10. It manages the write buffer 18. It manages the block erasure of FLASH memory 10. It manages the start-up and shut-down of the system after application and removal of power.

All sector writes from the host are directed to the write buffer 18. This minimises the effect of the slow write speed of FLASH memory. Sector data is transferred from the
10 write buffer 18 to the FLASH memory 10 as a background task. In applications where the write speed is not important the write buffer 18 may be omitted.

The system controller 16 is connected to a secondary memory 20 and to a host interface 22. The secondary memory 20 is used to store information such as defective block and erasable block tables for FLASH memory 10, address translation tables, error
15 logs, and microcontroller firmware. The secondary memory may be directly accessible from the host interface 22 to provide a directly mapped memory address space to the host. Referring to Figure 2, a main FLASH memory 10 is arranged in the following hierarchy of subdivisions. A FLASH memory consists of one or more semiconductor chips 30. A semiconductor chip 30 is a physical subdivision which does not have any logical
20 significance. An erase block 32 is a subdivision of a semiconductor chip 30. It is the minimum number of cells that may be deleted in an erase operation. The number of erase blocks in a semiconductor chip varies according to the size of the chip and the chip configuration. A rectangular group of word locations is assigned as a FLASH block. A

FLASH block 34 is the minimum maskable block size for defect management. If a cell is faulty then the minimum area which may be masked and replaced is a FLASH block 34. Defect mapping has a granularity of one FLASH block 34. A sector is one data unit according to disk emulation protocols. The data stored in FLASH memory (typically 512
5 bytes plus ECC bytes and a header) will hereinafter be referred to as a FLASH sector: the data sent by the host operating system (typically 512 bytes) will hereinafter be referred to as a host sector. A host sector is the basic unit of data transfer with the host. The most common host sector size is 512 bytes. A frame 36 is an addressable sub-unit of a FLASH block 34, for example a row within a rectangular FLASH block 34. Each word 38 is
10 randomly accessible as the minimum addressable unit at the technology independent interface 14. The word size is set by the width of the data I/O to the memory system: it may be one, two, or four bytes.

The actual physical location of bits within a word is transparent to the algorithm used in the present invention. The FLASH memory is treated simply as a rectangular array
15 of word locations.

The FLASH block 34 could be set to equal the host sector size (in most cases the host sector size is 512 bytes), the frame size or even the word size. The memory harvest (the fraction of gross memory that can be used for data storage) reaches a plateau below a certain value of FLASH block size, so that further reduction of the FLASH block size does
20 not result in much improvement in the memory harvest. There are, however, disadvantages associated with having a small FLASH block size. The smaller the FLASH block size the more storage management space is required. The component test complexity also increases. Thus, there is an optimum range for the size of a FLASH block.

Different addressing formats could be used with this invention without departing from the inventive concept.

For 100% memory utilisation the start address for a sector must be the word immediately following the end address of the preceding sector. To accomplish this the start address is defined to a resolution of a single word. Since the FLASH sectors are stored without any intervening gaps, FLASH sectors will overflow both FLASH blocks and erase blocks. In such a case a FLASH sector is known as a split sector and its continuation region is located at the start of the following FLASH block or erase block. The start address of FLASH sectors within FLASH blocks will therefore be different for each FLASH block. To ensure that this is easily manageable, it is necessary to ensure that the order in which erase blocks are written corresponds exactly to their physical addressing order. If there is a defective cell in a FLASH block then the block will be masked out. This may give rise to the situation where part of a FLASH sector is stored in one FLASH block, the next FLASH block is masked out, and the rest of the FLASH sector is stored in the FLASH block immediately above the one that was masked out. FLASH memory space may therefore be discontinuous.

Figures 3A and 3B illustrate how the write/erase algorithm manages the location and frequency of erase and write operations in main FLASH memory. Figure 3A shows a FLASH memory prior to an erase/write phase. Figure 3B shows the same FLASH memory after a write/erase phase.

The algorithm is based on the existence of two pointers, the write pointer and the erase pointer, defining current write and erase locations. The locations are defined as physical sector addresses and the pointers cycle through physical sector address space.

The write pointer defines the physical sector number to which the next sector write operation should be directed. The physical location of this sector is guaranteed to have been previously erased in preparation for writing data by the system controller 16 in the memory system. Sector data to be written may have originated from either the write buffer
5 18 following a sector write from the host or from another sector in main FLASH memory
10 in preparation for an erase block erase operation. Following a sector write operation, the write pointer is incremented by one to point to the next highest physical sector address. When the write pointer reaches the highest usable physical sector address it then wraps around and points at the lowest usable physical address.

10 The erase pointer defines the erase block address in FLASH memory space of the next erase block which is scheduled for erasure. A second field in the erase pointer defines the physical sector number of the first complete sector located at the bottom of the erase block. Following an erase block erase operation, the erase block number is incremented by one. A lookup table (hereinafter referred to as the erase block table) provides the
15 physical sector address of the first sector in the erase block that the erase pointer is pointing at. If the erase block is totally defective and contains no sectors which can be written to (signified by a specific entry in the erase block table) then the erase pointer is incremented again. When the erase pointer reaches the highest usable erase block it then goes back to the lowest part of the memory and points at the lowest usable erase block. Erase blocks are
20 erased at a rate which aims to maintain an approximately constant number of erased sector locations immediately above the current write location. An erase operation is therefore performed when the difference between the sector numbers defined in the erase pointer and write pointer falls below a certain threshold value. Prior to erasure of an erase block, a

sector containing valid data within that erase block must be relocated. Relocation of this valid data is accomplished by reading the sectors containing valid data into a relocate buffer (hereinafter referred to as a transfer buffer) and writing them back to the location pointed at by the write pointer. The logical to physical address translation table is updated to point
5 to the new physical sector address for the sector. Since only valid sectors are relocated in this way and since sectors with obsolete data are automatically erased, automatic compaction of the stored data is provided.

The process established by this algorithm can be considered equivalent to a small contiguous region of erased memory cycling through the entire FLASH memory space.
10 removing fragmented sectors it encounters at its upper boundary and compacting and relocating them at its lower boundary.

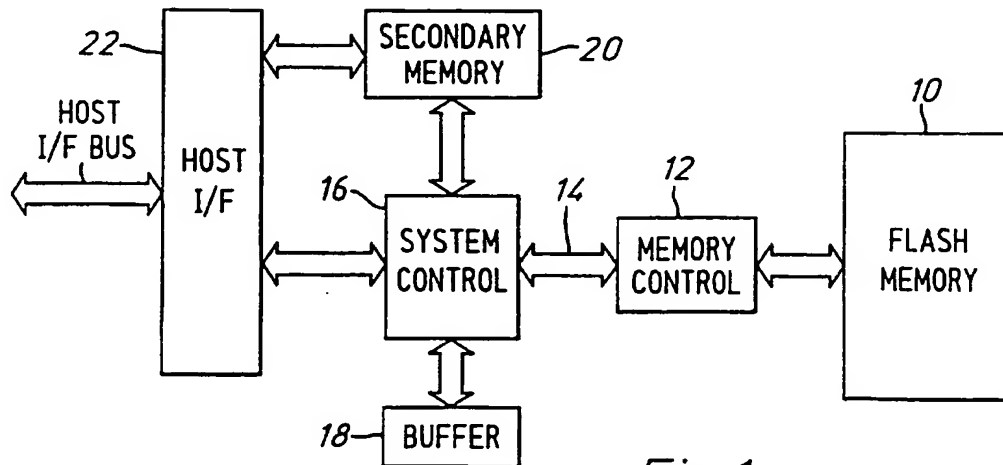
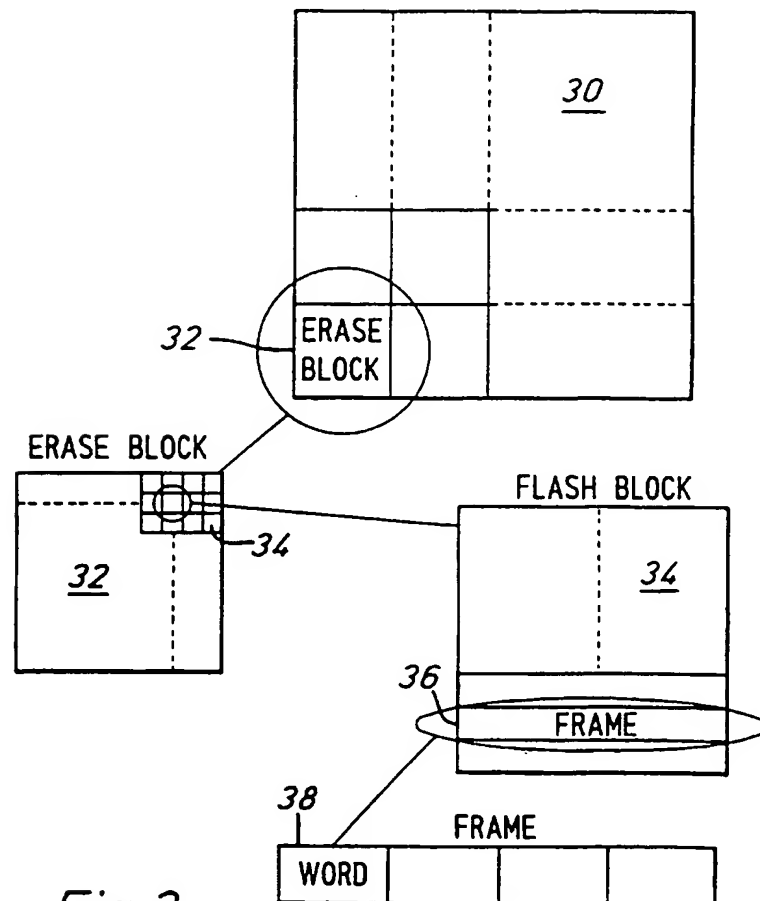
It will be appreciated that various modifications may be made to the above described embodiments within the scope of the present invention. For example. In an alternative arrangement, the contiguous region of erased memory cells moved in the other
15 direction: that is, if the lower boundary removed fragmented sectors which were immediately below the lower boundary and valid data was written to the location immediately below the upper boundary.

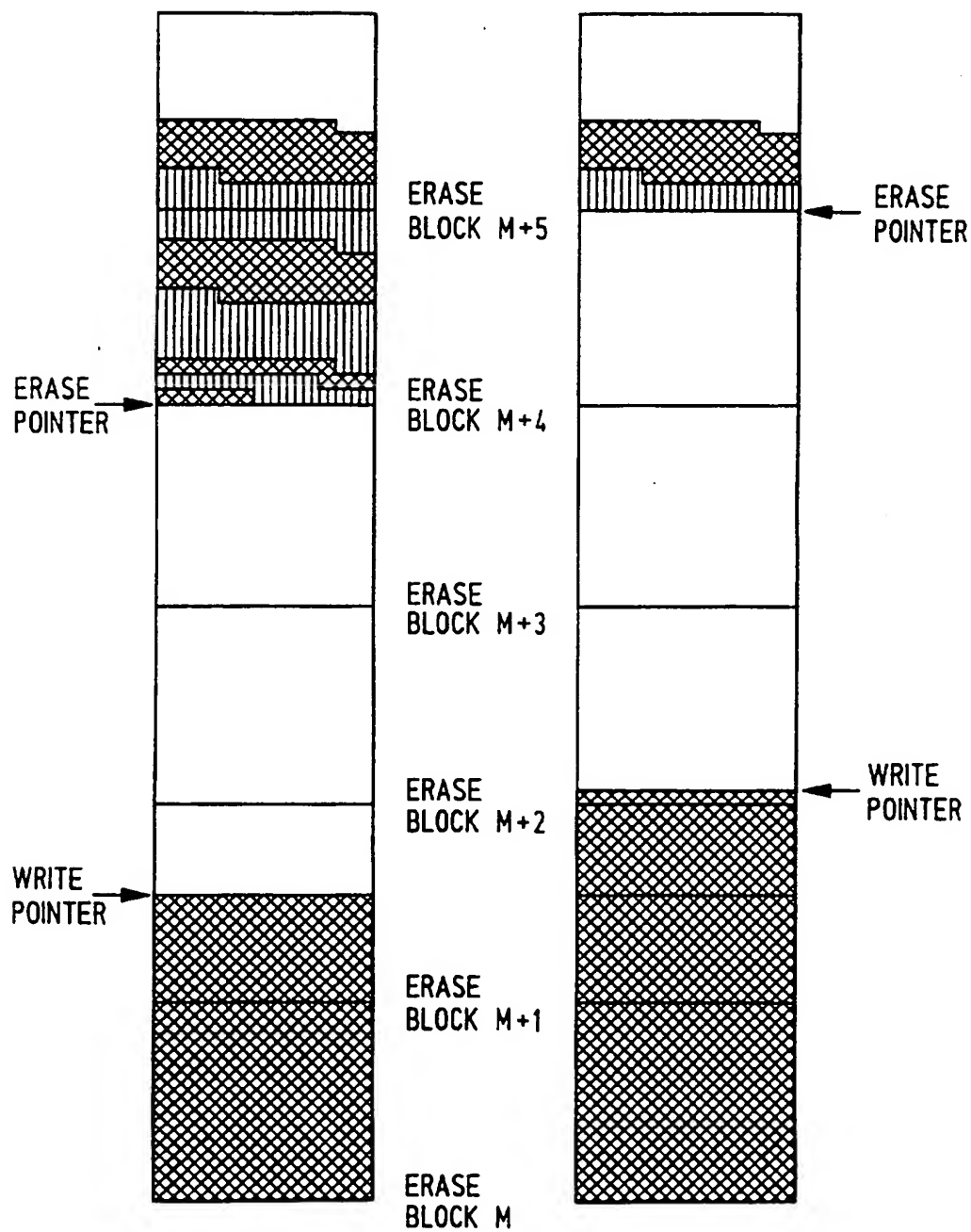
Claims

1. A memory arrangement comprising a plurality of storage means each having a pre-defined physical address. translation means for translating a logical address to a unique one of said pre-defined physical addresses, said storage means consisting of non-volatile
5 memory cells erasable in blocks; **characterised in that** said memory includes means for storing a first pointer to point to vacant storage means and means for storing a second pointer to point to the next unerased erasable block in a pre-determined sequence to the erasable block containing the said unwritten memory location; control means being provided to ensure that there is always at least one vacant erasable block between the
10 storage means indicated by the first pointer and the storage means indicated by the second pointer.
2. A solid state memory according to claim 1 **characterised in that** the unwritten location in the store is the lowest unwritten location in the store.
3. A solid state memory according to claim 1 **characterised in that** the unwritten
15 location in the store is the highest unwritten location in the store.
4. A solid state memory according to either of claims 1 or 2 **characterised in that** the next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is the unerased erasable block with the lowest physical address above the first pointer.
- 20 5. A solid state memory according to either of claims 1 or 3 **characterised in that** the next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is the unerased erasable block with the highest physical address below the first pointer.

6. A solid state memory according to any preceding claim **characterised in that** the next unerased erasable block in sequence to the erasable block containing the said unwritten memory location is determined by a suitable algorithm.
7. A solid state memory according to any preceding claim **characterised in that** a
5 buffer is used to increase the speed of writing data to the store.
8. A solid state memory according to claim 7 **characterised in that** said buffer is an SRAM buffer.
9. A solid state memory according to claim 7 **characterised in that** said buffer is a DRAM buffer.
- 10 10. A solid state memory according to any preceding claim **characterised in that** said store is FLASH EPROM.
11. A solid state memory according to any preceding claim **characterised in that** said the store is composed of chalcogenide cells.
12. A solid state memory according to any preceding claim **characterised in that** said
15 the store is a ferro-optic memory.
13. A solid state memory according to any preceding claim **characterised in that** said control means are provided by a microprocessor, a microcontroller, or a state machine.

1/2

*Fig. 1**Fig. 2*



INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 96/00250

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G06F12/02 G06F3/06 G11C16/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G06F G11C

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP,A,0 544 252 (FUJITSU LTD) 2 June 1993 see abstract; figure 1 see column 4, line 31 - line 36 see column 5, line 1 - line 10 see column 8, line 41 - column 9, line 35; figures 8,9	1,10
A	FR,A,2 665 791 (MAZINGUE DIDIER;REMERY PATRICK) 14 February 1992 see page 1, line 27 - page 2, line 28	1,10

☐ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *A* document member of the same patent family

Date of the actual completion of the international search

20 May 1996

Date of mailing of the international search report

17.06.96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+ 31-70) 340-3016

Authorized officer

Ledrut, P

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 96/00250

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0544252	02-06-93	JP-A- 5151097	18-06-93
		EP-A- 0686976	13-12-95
FR-A-2665791	14-02-92	NONE	